

## AMENDMENTS TO THE CLAIMS

5. (Previously Amended) A videoconferencing system comprising:  
a conference bridge for interconnecting a plurality of remotely located videoconferencing stations; and  
means for determining whether a conferee is speaking based, at least in part, on whether visual lip movements of said conferee are reasonably consistent with an audio signal from a conference station in which said conferee is located.
6. (Previously Added) The videoconference system of claim 5 wherein said means for determining whether said conferee is speaking comprises a voice activity detector.
7. (Previously Added) The videoconference system of claim 6 wherein said voice activity detector is implemented at each of said conference stations.
8. (Previously Added) The videoconference system of claim 6 wherein said voice activity detector is implemented at each of said conference bridge.
9. (Previously Added) The videoconference system of claim 6 wherein said voice activity detector includes image analysis and recognition software.

10. (Previously Amended) The videoconference system of claim 29 wherein said means for visually altering said image comprises means for highlighting a border around said image of said conferee determined to be speaking.

11. (Previously Amended) A videoconference station comprising:

a transmitter to transmit a combined audio video signal to a videoconference bridge; and

means for determining whether a conferee located at said videoconference station is speaking by analyzing whether visual lip movements of said conferee are substantially consistent with an audio signal at said station so as to indicate human speech.

12. (Previously Added) The videoconference station of claim 11 wherein said means for determining whether said conferee is speaking is a voice activity detector.

13. (Previously Added) The videoconference system of claim 12 wherein said voice activity detector includes image analysis and recognition software.

14. (Previously Amended) A method of displaying images of a plurality of conferees in a videoconference system, comprising:

determining whether a conferee is speaking by analyzing a consistency between visual lip movements of said conferee and an audio signal from a conference station in which said conferee is located such that the combination of lip movement and audio signal indicates human speech; and

visually altering an image of said conferee that is displayed to other conferees when said conferee is determined to be speaking.

15. (Previously Amended) A method of determining whether a conferee in a videoconference is speaking, comprising analyzing whether visual lip movements of said conferee are reasonably consistent with an audio signal from a conference station in which said conferee is located such that the combination of lip movement and audio signal indicates human speech.

16. (Previously Added) A method of determining whether a conferee in a videoconference is speaking comprising:

detecting that the conferee's lips are moving based on a visual image of the conferee's lips;

detecting an audio signal associated with the conferee; and

determining whether the conferee is speaking based, at least in part, on whether the detected audio signal is substantially consistent with the detected movement of the conferee's lips.

17. (Previously Added) The method of claim 16 further comprising:

altering an image of the conferee that is displayed to other conferees if the conferee is determined to be speaking.

18. (Previously Added) The method of claim 16 further comprising:

providing textual information to identify the conferee to other conferees if the conferee is determined to be speaking.

19. (Withdrawn) A method for identifying which conferee in a video conference is speaking comprising:

- detecting a first audio signal associated with a first conferee;
- detecting a second audio signal associated with a second conferee;
- comparing the first detected audio signal with the second detected audio signal to determine which detected audio signal is louder; and
- providing an indication to the first conferee and the second conferee of which detected audio signal is louder.

20. (Withdrawn) The method of claim 19, wherein providing an indication to the first conferee and the second conferee of which detected audio signal is louder comprises:

- altering an image of either the first conferee or the second conferee based, at least in part, on which detected audio signal is determined to be louder.

21. (Withdrawn) The method of claim 20, wherein altering the image of either the first conferee or the second conferee based, at least in part, on which detected audio signal is determined to be louder comprises:

- highlighting a border around the image of either the first conferee or the second conferee based, at least in part, on which received audio signal is determined to be louder.

22. (Withdrawn) The method of claim 19, wherein providing an indication to the first conferee and the second conferee of which detected audio signal is louder comprises:

providing textual information to the first conferee and the second conferee indicating which detected audio signal is determined to be louder.

23. (Previously Added) A videoconferencing system comprising:

a conference bridge for interconnecting a plurality of remotely located videoconferencing stations; and

a voice activity detector to determine whether a conferee is speaking based, at least in part, on whether lip movements of the conferee are substantially consistent with an audio signal from a conference station in which said conferee is located.

24. (Previously Added) The videoconference system of claim 23, wherein the voice activity detector is implemented at each of the plurality of videoconference stations.

25. (Previously Added) The videoconference system of claim 23, wherein the voice activity detector is implemented at the conference bridge.

26. (Previously Added) A videoconference station comprising:

a transmitter to transmit a combined audio/video signal to a videoconference bridge; and

a voice activity detector to analyze whether visual lip movements of a conferee are substantially consistent with an audio signal at the videoconference station so as to indicate human speech.

27. (Previously Added) The videoconference station of claim 26, wherein the voice activity detector includes image analysis and recognition software.

28. (Previously Added) The videoconference station of claim 26, further comprising:  
a display unit to provide a visual representation of conferees participating in a videoconference.

29. (Previously Added) The videoconferencing system of claim 5, further comprising:  
means for visually altering an image of said conferee displayed in other conference stations if said conferee is determined to be speaking.